Comparison of the performance of several recent hydrogen combustion mechanisms

CrossMark

Carsten Olm [a,b], István Gy. Zsély [a], Róbert Pálvölgyi [a], Tamás Varga [a,b], Tibor Nagy [a], Henry J. Curran [c], Tamás Turányi [a,*]

[a] *Institute of Chemistry, Eötvös University (ELTE), Budapest, Hungary*
[b] *MTA-ELTE Research Group on Complex Chemical Systems, Budapest, Hungary*
[c] *Combustion Chemistry Centre, National University of Ireland, Galway (NUIG), Ireland*

ARTICLE INFO

ABSTRACT

A large set of experimental data was accumulated for hydrogen combustion: ignition measurements in shock tubes (770 data points in 53 datasets) and rapid compression machines (229/20), concentration–time profiles in flow reactors (389/17), outlet concentrations in jet-stirred reactors (152/9) and flame velocity measurements (631/73) covering wide ranges of temperature, pressure and equivalence ratio. The performance of 19 recently published hydrogen combustion mechanisms was tested against these experimental data, and the dependence of accuracy on the types of experiment and the experimental conditions was investigated. The best mechanism for the reproduction of ignition delay times and flame velocities is Kéromnès-2013, while jet-stirred reactor (JSR) experiments and flow reactor profiles are reproduced best by GRI3.0-1999 and Starik-2009, respectively. According to the reproduction of all experimental data, the Kéromnès-2013 mechanism is currently the best, but the mechanisms NUIG-NGM-2010, ÓConaire-2004, Konnov-2008 and Li-2007 have similarly good overall performances. Several clear trends were found when the performance of the best mechanisms was investigated in various categories of experimental data. Low-temperature ignition delay times measured in shock tubes (below 1000 K) and in RCMs (below 960 K) could not be well-predicted. The accuracy of the reproduction of an ignition delay time did not change significantly with pressure and equivalence ratio. Measured H_2 and O_2 concentrations in JSRs could be better reproduced than the corresponding H_2O profiles. Large differences were found between the mechanisms in their capability to predict flow reactor data. The reproduction of the measured laminar flame velocities improved with increasing pressure and total diluent concentration, and with decreasing equivalence ratio. Reproduction of the flame velocities measured using the flame cone method, the outwardly propagating spherical flame method, the counterflow twin-flame technique, and the heat flux burner method improved in this order. Flame cone method data were especially poorly reproduced. The investigation of the correlation of the simulation results revealed similarities of mechanisms that were published by the same research groups. Also, simulation results calculated by the best-performing mechanisms are more strongly correlated with each other than those of the weakly performing ones, indicating a convergence of mechanism development. An analysis of sensitivity coefficients was carried out to identify reactions and ranges of conditions that require more attention in future development of hydrogen combustion models. The influence of poorly reproduced experiments on the overall performance was also investigated.

© 2014 The Combustion Institute. Published by Elsevier Inc. All rights reserved.

1. Introduction

The elementary reactions of the combustion of hydrogen are a central part of the mechanisms which describe the combustion of all hydrocarbon and oxygenated hydrocarbon fuels. Moreover,

hydrogen is an important fuel in itself in areas like the carbon-free economy, in fuel safety issues, and for rocket propulsion. In accordance with its high significance, several new hydrogen combustion mechanisms were published in the last decade. In these publications, agreement between measurements and simulations is usually characterized by plots, in which the experimental data and the simulation results are depicted together. However, quantitative agreement of the simulation results with the experimental data has not been investigated. A quantitative evaluation allows for

the determination of experiments that are well estimated by simulations in contrast to those that are insufficiently described. Furthermore, strengths and weaknesses of the mechanisms in certain ranges of operating conditions can be detected. Knowing the specific behavior of a mechanism helps to reduce uncertainties in the description of experiments during mechanism development and optimization.

This paper has several novelties compared to the previous publications containing comparisons of combustion mechanisms. The comparison is performed on a very comprehensive set of experiments; various measurement types (ignition delay time, flow reactor, JSR and flame velocity measurements) and experimental techniques (e.g., shock tube and RCM experiments) are included in the analysis. All important hydrogen reaction mechanisms published in the last decade are considered, as well as syngas and hydrocarbon oxidation mechanisms that were previously used to describe hydrogen reactions. The performance of all of these reaction mechanisms is compared in detail, and the conclusions drawn are supported by objective numbers.

2. Methodology

In this work the agreement of experimental and simulation results is investigated using the following objective function

$$E_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \left(\frac{Y_{ij}^{sim} - Y_{ij}^{exp}}{\sigma\left(Y_{ij}^{exp}\right)} \right)^2$$

$$E = \frac{1}{N} \sum_{i=1}^{N} E_i$$

Where

$$Y_{ij} = \begin{cases} y_{ij} & \text{if } \sigma(y_{ij}^{exp}) \approx \text{constant} \\ \ln y_{ij} & \text{if } \sigma(\ln y_{ij}^{exp}) \approx \text{constant} \end{cases}$$

Here N is the number of datasets and N_i is the number of data points in the ith dataset. Values y_{ij}^{exp} and $\sigma(y_{ij}^{exp})$ are the jth data point and its standard deviation, respectively, in the ith dataset. The corresponding simulated (modeled) value is Y_{ij}^{sim} obtained from a simulation using an appropriate detailed mechanism. If a measured value is characterized by absolute errors (the scatter can be considered independent of the magnitude of y_{ij}), then $Y_{ij} = y_{ij}$. We used this option for laminar flame velocities and measured concentrations. If the experimental results are described by relative errors (the scatter is proportional to the value of y_{ij}), then we used the option $Y_{ij} = \ln y_{ij}$, which is characteristic for ignition time measurements. Error function values E_i and E are expected to be near to unity if the chemical kinetic model is accurate, and deviations of the measured and simulated results are caused by the scatter of the experimental data only. Note that due to the squaring in the definition of E, a twice as high deviation of the simulated and experimental values of one mechanism in comparison to another leads to a four times higher value of E. This objective function has been used in our previous studies on the estimation of rate parameters from experimental data [1–3].

In addition to the average error function E, another quantity was used to analyze the behavior of the mechanisms. The average absolute deviation D is defined with the absolute deviation of an individual data point D_{ij} as:

$$D_{ij} = Y_{ij}^{sim} - Y_{ij}^{exp}$$

$$D = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N_i} \sum_{i=1}^{N_i} D_{ij}$$

using the same transformation as mentioned above. In contrast to E, the sign of the difference $Y_{ij}^{sim} - Y_{ij}^{exp}$ is maintained in the definition of D. Trends such as systematic under- or over-prediction are thereby captured in the D_{ij} values.

It is possible to characterize the similarity of simulation results using different mechanisms by calculating correlation coefficients based on the values of D_{ij}. Similar to the definitions of E and D, correlation coefficients C are calculated for each dataset and then averaged over all of the N datasets (if more than one concentration was measured in an experiment, e.g., in a flow reactor or JSR, correlation values are first determined for each species separately and then averaged). In this averaging step, datasets with $N_i \leqslant 2$ were not considered as they would perturb the average correlation values with unrealistic values (-1 or $+1$). Correlation of the individual error function values would not provide the same meaningful information about the similarity of two mechanisms, since the sign of the deviation is lost due to squaring. Hence, a positive and a negative deviation of exactly the same extent would give a correlation value of $C = 1$, which is misleading. Also, the average deviation D is not suitable for the comparison of the agreement between the experimental and simulation results using different mechanisms, since at the two summations of the D_{ij} values large positive and negative deviations might even out.

3. Mechanisms

Our aim was to test all major hydrogen combustion mechanisms that were published since 1999. Table 1 contains the list of the investigated mechanisms. Multiple mechanisms from the same research group were tested only if the older mechanism is conceptually different from the newer one. Otherwise, only the latest mechanism was considered. For instance, three mechanisms published by the Galway group were included in the mechanism comparison. Hereby, the hydrogen core of the NUIG-NGM-2010 mechanism is an extension of the previous model (ÓConaire-2004) since the bath gas helium was added. Furthermore, some of the reaction rates differ substantially between the three mechanisms, e.g., for the reaction H + O2 = O + OH. Given the importance of this reaction, these changes result in significant differences in the simulation of the hydrogen combustion chemistry. Further differences are detailed in [4]. The latest mechanism of this group, Kéromnès-2013, also features a sub-model for the reactions of the excited OH radical (OH*) largely based on the work of [5] and updated by [6]. These reactions have shown to be of high importance for a more accurate reproduction of shock tube ignition studies at high temperatures (e.g., those reported in [7]).

The hydrogen subset of reactions in the Li-2007 mechanism is identical to their hydrogen mechanism published in 2004 [8], however, the 2007 version seems to be cited more often in the recent literature, which is the reason why this name was used in the present work. In the further references to the mechanisms, an identifier (as listed in Table 1) is used, which combines the name of the first author(s) or the research group, and the year of publication.

Some of these mechanisms were originally developed for the description of hydrogen combustion [9–12], but we investigated also other mechanisms that were developed for syngas combustion [4,13–19], or the combustion of hydrocarbons or oxygenates [20–26]. For the latter mechanisms, Table 1 refers to the number of species and reactions of the hydrogen combustion part of these mechanisms; values in brackets define the size of the original mechanisms. The program MECHMOD [27] was used to remove unnecessary species and reactions describing the carbon chemistry from the mechanisms.

All mechanisms are able to describe the combustion of hydrogen–air mixtures and thus are able to handle N2 as a bath gas.

Table 1

The hydrogen combustion mechanisms investigated, and the number of species and reactions in these reaction mechanisms. The numbers in parenthesis indicate the corresponding figures in the original mechanisms. All mechanisms can handle N_2 bath gas and the table indicates if the mechanism can also cope with Ar and He bath gases. The values of the average error function for all mechanisms are given for six cases: A – Ignition delay times, all diluents (875 data points/73 datasets), B – JSR concentrations, all profiles (49/9), C – Flow reactor concentrations, all profiles (191/14), D – Flame velocities, all diluents except He (420/49), E – Overall results, all diluents except He (1529/144), F – Overall results, all diluents including He (1652/168). Overall results are indicated by bold letters.

No.	Mechanism ID	Ref.	Species number (orig.)	Ar/He	Reactions number (orig.)	Average error function values					
						A	B	C	D	E	F
1	Kéromnès-2013	[4]	12 (17)	×/×	33 (49)	11.9	3.0	13.3	13.9	**12.2**	**12.4**
2	NUIG-NGM-2010	[20]	11 (293)	×/×	21 (1593)	14.0	3.0	7.3	20.2	**14.8**	**13.3**
3	ÓConaire-2004	[9]	10	×/–	21	15.4	3.0	8.2	18.5	**15.0**	–
4	Konnov-2008	[10]	10	×/–	33	19.7	3.1	10.9	15.2	**16.3**	–
5	Li-2007	[21]	11 (21)	×/×	25 (93)	20.7	3.0	7.8	16.0	**16.8**	**15.8**
6	Hong-2011	[12]	10	×/–	31	14.5	3.0	8.1	28.5	**17.9**	–
7	Burke-2012	[11]	11	×/×	27	26.6	3.1	3.9	14.6	**18.9**	**16.7**
8	SaxenaWilliams-2006	[13]	11 (14)	×/×	21 (30)	23.8	3.0	28.3	16.5	**20.5**	**19.0**
9	Davis-2005	[14]	11 (14)	×/×	25 (38)	36.7	3.0	4.9	16.4	**24.7**	**21.8**
10	Starik-2009	[15]	12 (16)	×/×	26 (44)	37.4	3.4	3.8	24.4	**27.7**	**24.7**
11	USC-II-2007	[22]	10 (111)	×/–	28 (784)	36.2	3.0	4.7	26.1	**27.7**	–
12	CRECK-2012	[16]	11 (14)	×/×	21 (34)	15.2	2.9	21.4	56.9	**29.1**	**25.6**
13	SanDiego-2011	[23]	11 (50)	×/×	21 (244)	78.0	3.0	27.7	16.5	**47.7**	**42.0**
14	GRI3.0-1999	[24]	10 (53)	×/–	29 (325)	71.4	2.4	11.6	32.0	**48.0**	–
15	Sun-2007	[17]	11 (15)	×/×	32 (48)	97.9	3.1	25.4	26.7	**61.0**	**53.8**
16	Rasmussen-2008	[18]	10 (24)	×/–	30 (105)	197.1	3.0	17.8	35.4	**113.1**	–
17	Ahmed-2007	[25]	10 (246)	×/–	20 (1284)	257.9	3.1	3.9	20.7	**137.4**	–
18	Zsély-2005	[19]	10 (13)	×/–	32 (44)	544.3	3.2	15.6	26.0	**284.5**	–
19	Dagaut-2003	[26]	9 (132)	–/–	21 (922)	–	3.1	4.9	–	–	–

Table 1 indicates if the mechanism can also be used with Ar and/or He bath gases. The mechanism of Dagaut-2003 [26] includes neither Ar nor He as possible bath gases. This affects the comparison with respect to ignition delay times and flame velocities, where either Ar or He was used as a bath gas. A direct comparison of the Dagaut-2003 mechanism to the other ones is possible only in experiments where neither Ar nor He were used. For this reason, the Dagaut-2003 mechanism was handled separately in all ignition delay time and flame velocity comparisons (see Sections 6.1 and 6.4).

The thermochemical and transport data were used as published online and/or provided by the authors directly. Transport data were not provided with the Rasmussen-2008 mechanism [18] as it was tested by simulating only shock tube and flow reactor measurements. However, the transport coefficients of all species involved in hydrogen combustion were identical in all investigated mechanisms. For this reason, the same transport data were also used for the Rasmussen-2008 mechanism. In addition to the commonly used set of species, O_3 was defined in Starik-2009 as well as the aforementioned OH^* in the Kéromnès-2013 mechanism. The transport data provided by the respective authors were used.

The numbering of the mechanisms in Table 1 is according to their overall performance from the best (1) to the worst (18). The Dagaut-2003 mechanism, which could not be tested over the complete set of experimental data, was given the number 19.

4. Experimental data

A large set of experimental data related to the combustion of hydrogen was collected. These types of measurements, called indirect measurements or bulk measurements, have been used to test detailed reaction mechanisms. We utilized all measurements that were used for testing the recent mechanisms ÓConaire-2004 [9], Konnov-2008 [10], Hong-2011 [12], Burke-2012 [11], and Kéromnès-2013 [4]. References for measurements were collected from these recent review articles, and the experimental data were digitized from the original publications. Also, a comprehensive search was carried out to find all measurements that can be used to test hydrogen combustion mechanisms. Our data collection is therefore much wider-ranging than any other set of data that has

ever been used for the development of hydrogen combustion mechanisms. The detailed list of the data with references is given in the Supplementary material (Part 1).

The data include ignition measurements in shock tubes (770 data points in 53 datasets from 16 original publications) and rapid compression machines (229 data points in 20 datasets from 3 publications), species concentrations in flow reactors (191 utilized data points (389 total) in 14 datasets (17 total) from 2 publications) and jet-stirred reactors (149 data points in 9 datasets from one publication) and flame velocity measurements (543 utilized data points (631 total) in 73 datasets from 22 publications). A dataset contains those data points that were consecutively measured on the same apparatus at similar conditions except for one factor that was systematically changed. Usually, publications on measurements contain one or only a few datasets.

It is unavoidable that some regions of operating conditions are over-represented in the database due to the fact that they were carried out multiple times by various groups of researchers, e.g., to validate their experimental apparatuses. This introduces a sampling bias that could be mitigated by weighing down or averaging those experimental results. However, for the present study we decided to treat all measured data points with equal weights, even those measurements that were carried out at very similar conditions.

All relevant experimental conditions and results for each dataset were encoded in the PrIMe file format [28], which is an XML scheme used for the systematic storage of various kinds of combustion experiments. These stored XML data provide all the information required for the simulation of experiments and the calculation of properties observed or derived from experiments (e.g., the ignition delay as defined in the corresponding experimental publication). A MATLAB code was written which allows automatic CHEMKIN-II [29] simulations and error evaluations for a specific mechanism on the full data set. The MATLAB code starts the corresponding simulation code, collects the simulation results, and evaluates the deviation between the simulated and the experimental results. In principle, the complete investigation of a mechanism against several thousand experimental data can be carried out in a single run. The chemical mechanism was then replaced with another mechanism and the same procedure was repeated.

5. Simulation of the experiments

The relationship between the types of experiments, modeling approaches and computational codes to solve the respective problem is shown in Fig. 1.

5.1. Shock tube simulations

Shock tube experiments were simulated using the SENKIN program [30] of the CHEMKIN-II package. Constant volume and adiabatic conditions were assumed. Ignition delay times were extracted from the simulated pressure and concentration profiles as described in the respective publications, e.g., based on the maximum slope of the pressure profile. If an ignition criterion could not be modeled, another similar criterion was chosen. For example, the maximum concentration of the excited radical OH^* as measured in [7] cannot be modeled with a mechanism that does not contain this species. For all mechanisms except Kéromnès-2013, the criterion "OH maximum" had to be used instead, which has shown to give similarly good predictions except at very high temperatures behind the reflected shock wave.

Low-temperature shock tube experiments, where the ignition delays are in the order of milliseconds, should be handled with special care by considering the possibility of a pressure change during the induction period [31]. At such reaction times the pressure behind the reflected shock wave increases with time [32]. In both articles it is demonstrated that by taking into account this facility effect the overall description of the experiments by the models improves. It has also been discussed by Dryer and Chaos [33,34] that at these conditions the measured ignition delays are extremely sensitive to impurities in the mixture. They have shown that by assuming a catalytic conversion between H_2O_2 and OH, the description of the experiments also improves.

Unfortunately the measurements we collected did not have pressure profiles published along with the measured ignition delay times; hence, we could not take into account such facility effects. Alternatively, data points where long ignition delay times above 1–2 ms were measured and/or data points at low temperatures can be excluded. The latter option was chosen in the present study.

5.2. RCM simulations

RCM experiments were simulated using the VTIM option of SENKIN [30], to account for the effects of compression and heat loss. Volume–time histories were calculated from the experimental pressure profiles in the case of the measurements of Kéromnès

et al. [4] and Das et al. [35]. Assuming an adiabatic reaction core [35], the experimentally recorded pressure profiles were converted to the corresponding volume profile which was used for the simulations. The measurements of Mittal et al. [36] were published together with volume–time histories expressed in polynomial form. The simulations were performed using these volume–time histories. Some of their experiments lacked published heat loss data and these had to be excluded from the comparison.

5.3. Flow reactor simulations

Flow reactor experiments were performed using SENKIN [30], assuming constant volume and adiabatic conditions. The half depletion of the measured concentration of the fuel was matched to the simulated concentration profile to take into account the time shift due to mixing effects. Only those flow reactor measurement points were used where the concentration of the consumed fuel (hydrogen) is between 10% and 90% of the initial fuel concentration.

5.4. Jet-stirred reactor simulations

The simulations were performed using the PSR code [37] of the CHEMKIN-II package. Wall effects are expected to play a dominant role at low temperatures. The corresponding data points were excluded from the comparisons in the present study.

5.5. Laminar flame velocity simulations

Laminar flames were simulated using the PREMIX code [38] of the CHEMKIN-II package. Thermo-diffusion (Soret effect) was taken into account, and molecular diffusion was described with the multicomponent diffusion approach. The number of the grid points was always at least 600 to minimize the effect of the grid size on the simulated laminar flame velocity. This lower threshold was determined in a preliminary grid dependence study.

6. Results and discussion

The results of the testing are sorted according to the types of experiments. In general, five types of figures are presented: (a) a plot showing the overall performance of the mechanisms vs. their year of publication (Fig. 16), (b) correlation matrices using a characteristic coloring (e.g., Fig. 2), (c) detailed investigation of the influence of the operating conditions (e.g., Fig. 5), (d) Column plots illustrating the influence of the diluents and the experimental
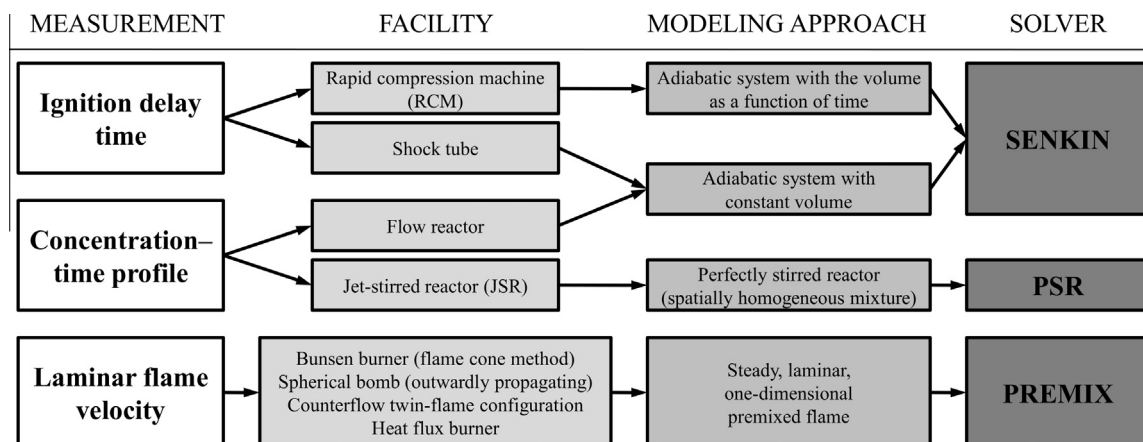

Fig. 1. Overview of the types of measurements, facilities, modeling approaches and the appropriate solver from the CHEMKIN-II package for the given problem.
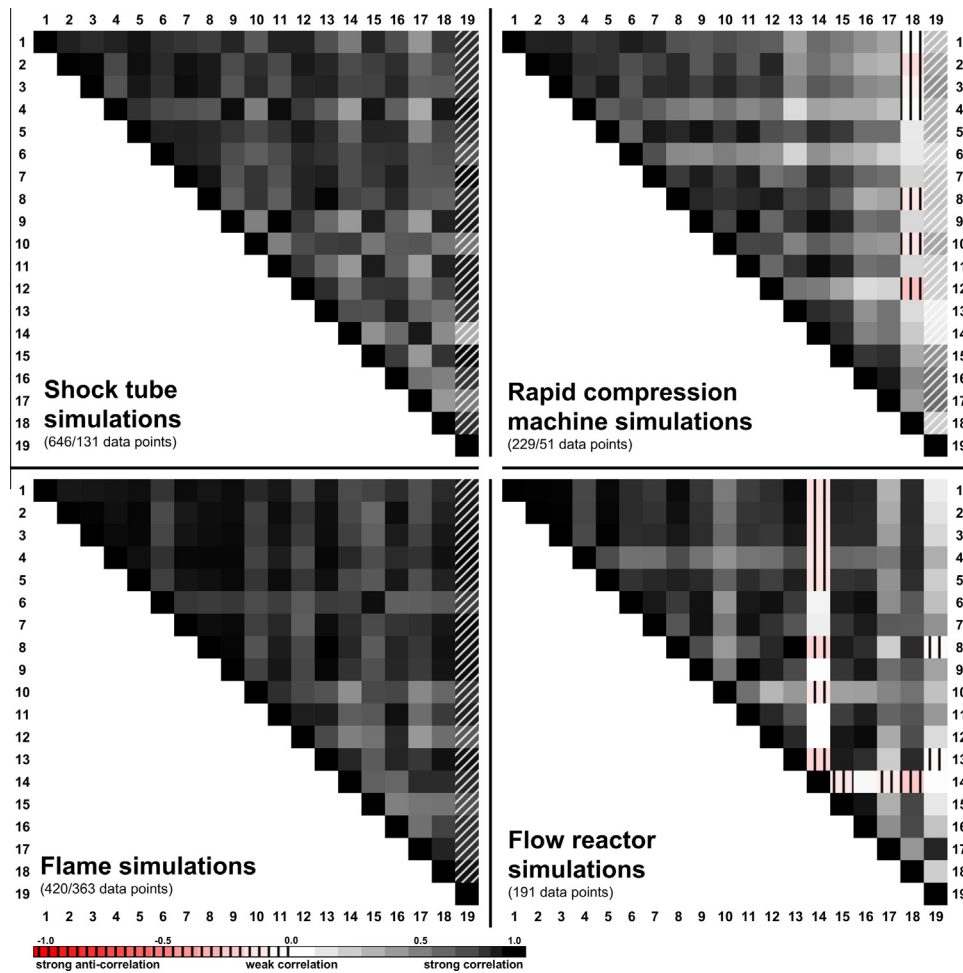
Fig. 2. Matrix of correlations of absolute deviation values for shock tube simulations (top left), rapid compression machine simulations (top right), flame simulations (bottom left) and flow reactor simulations (bottom right). Diagonal stripes were used when the comparison was based on a different dataset (for Dagaut-2003: Ar and He excluded). Not shown: correlations of jet-stirred reactor simulations (see text).
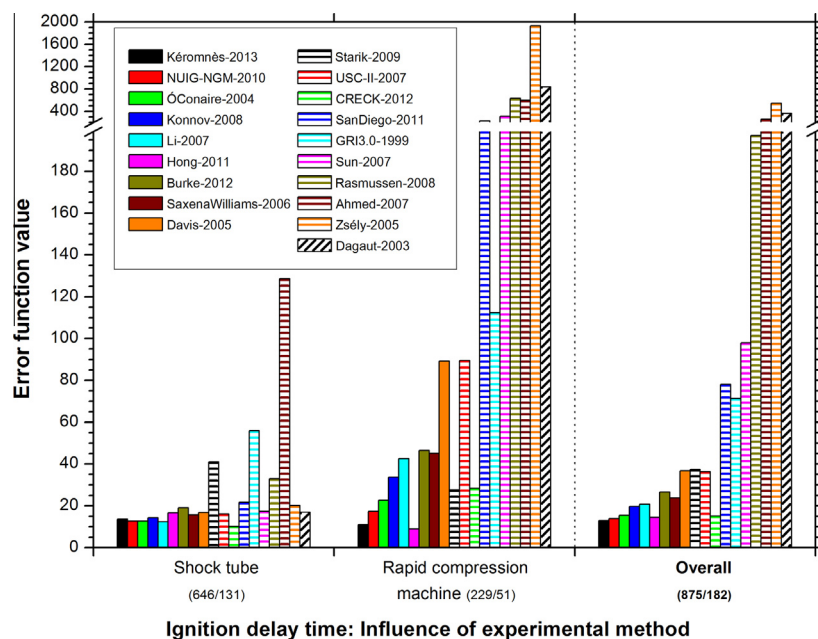

Fig. 3. Error of the reproduction of the ignition delay times according to the type of measurement. The numbers in parentheses indicate the number of data points included. The second values refer to the number of data points for Dagaut-2003.

methods (e.g., Fig. 3) and (e) a plot comparing the average absolute deviation values for all mechanisms (Fig. 12). Figure types (a), (c) and (d) are based on the E and E_i values, while types (b) and (e) on values of D and D_{ij}.

In Fig. 2, the correlations of the mechanisms with respect to the absolute deviation values (simulation *vs.* measurement) are shown. Mechanisms are referred to with their respective identifier (see Table 1). The correlations will be discussed and compared in the respective chapters.

6.1. Ignition delay times

Figure 3 shows that about half of the investigated mechanisms provide a similarly good reproduction of the measured ignition delay times (right group of columns in Fig. 3). The worst mechanisms in this respect are Zsély-2005, Ahmed-2007, Rasmussen-2008 and Dagaut-2003 (experiments N_2 and N_2/H_2O only), mainly due to their inability to reproduce RCM experiments (middle group). For shock tube data (left group), only Ahmed-2007 stands out in a negative sense. Note that the Zsély-2005 mechanism [19] contains the rate parameters recommended by Baulch et al. [39] without modification. This shows that although the evaluated rate parameter values of Baulch et al. are widely used in the creation of combustion mechanisms, further tuning is needed for a good description of ignition delay times in the hydrogen combustion system.

Figure 4 shows that there is no clear dependence of the accuracy of predictions on the bath gas or diluent mixture that was used in shock tube experiments, even though some mechanisms perform significantly worse for certain diluents (e.g., Starik-2009 for N_2-based diluents or Ahmed-2007 for Ar). This gives an indication that in some of the mechanisms, third body collision coefficients might not be chosen ideally in order to match experimental data. Among the nine best overall mechanisms (full columns), only Konnov-2008 and Hong-2011 perform better for RCM data measured in N_2 or N_2/H_2O [35,36], all others perform better for the measurements using Ar/N_2 [4].

Figure 5 shows the performance of the mechanisms compared to ignition delay measurements in shock tubes and RCMs according to interval ranges of temperature, pressure, equivalence ratio and diluent concentration, respectively. The results for Dagaut-2003 are based on a different subset of data (no Ar dilution). To

visually distinguish it from the other mechanisms, the symbols corresponding to this mechanism are indicated with triangles in line plots and with diagonal stripes in column plots. As for all partial comparisons, intervals were defined based on a careful inspection of the available data taking into account the specifics of the respective type of measurement and ensuring a statistically significant number of data points in each interval at the same time. For each interval, two numeric values referring to the number of data points considered are given. The second value refers to the number of data points considered for Dagaut-2003, while for all other mechanisms the first value is applicable. The shock tube results of Fig. 5 will be discussed first, followed by a discussion of the RCM results.

In all shock-tube related plots, Ahmed-2007 can be identified as the worst performing one and will not be discussed in more detail. As it can be seen in Fig. 5 (top left panel, left part), most mechanisms reproduce the shock tube experiments similarly well at temperatures above 1000 K, while below 1000 K the reproduction of the measured ignition delay times is generally poor. This behavior could be associated with the facility effect of the shock tube experiments as described in Section 5.1. For the purpose of an overall mechanism comparison, shock tube ignition delay data measured at $T \leqslant 1000$ K were excluded. It is worth stressing that this lower threshold is not a strict limit and it does not filter out all data points with long measured ignition delays. However, it has been shown previously [33,34] that at these conditions, the discrepancies between experimental results and homogeneous adiabatic simulations increase significantly.

As Fig. 5 (bottom left panel, left part) indicates, simulation errors from most mechanisms do not show a distinct pressure dependence, except Rasmussen-2008 and Dagaut-2003, which reproduce the experimental data less accurately at higher pressures. Interestingly, all of the mechanisms, barring Dagaut-2003 (for which results are based on a different data subset) and GRI3.0-1999, perform very similarly well in the intermediate pressure interval of 1.17–3.0 atm and at near-atmospheric conditions.

Clear trends cannot be observed with respect to the dependence on the equivalence ratio in shock tube experiments (Fig. 5, top right panel, left part). However, the majority of the mechanisms (expect for Rasmussen-2008 and Zsély-2005) show better performance at higher equivalence ratios.
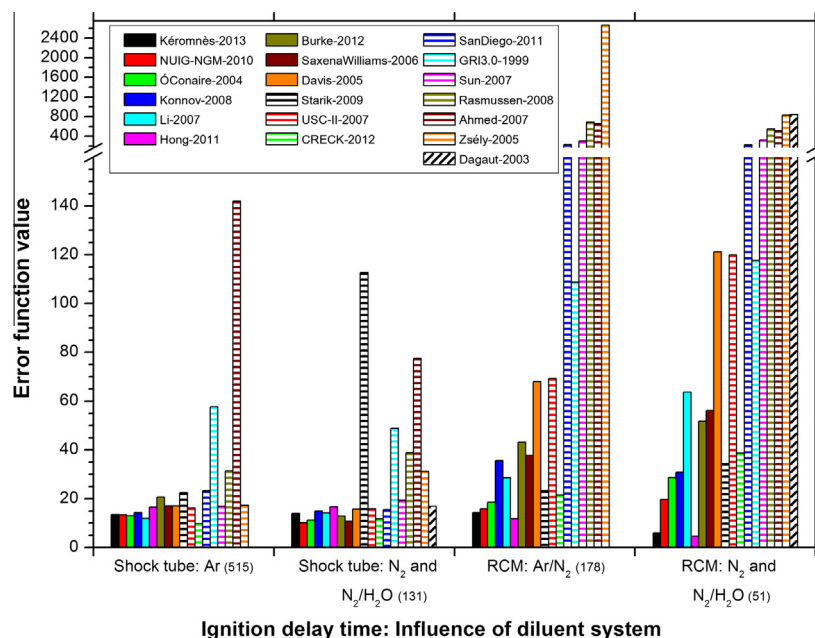

Fig. 4. Performance of the mechanisms with respect to ignition delay measurements in shock tubes and rapid compression machines for various diluent systems.
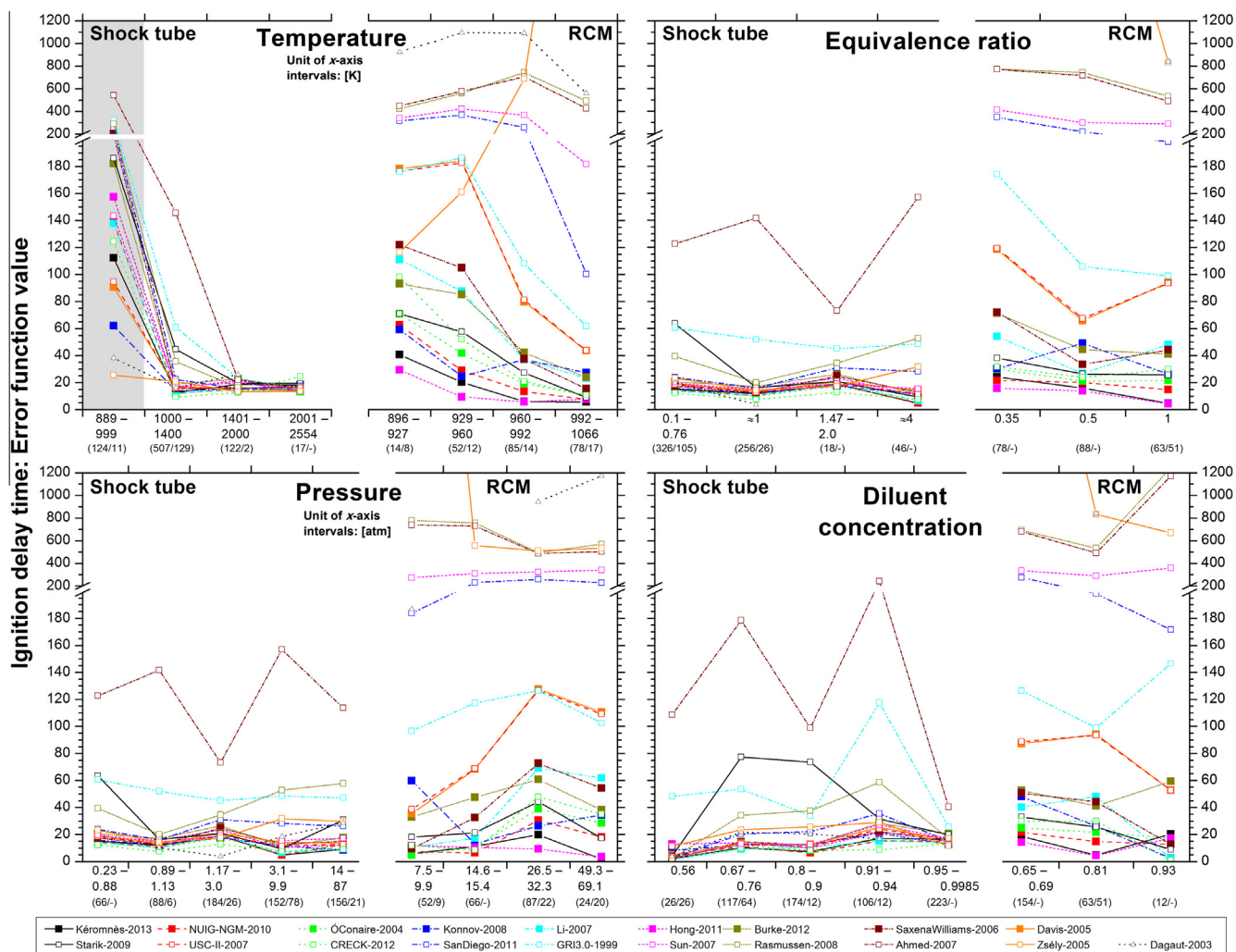
Fig. 5. Performance of the mechanisms for various ranges of temperature, pressure and equivalence ratio and diluent concentration with respect to ignition delay time. Each plot shows the results for shock tubes (left part) and rapid compression machines (right part). Gray shaded: shock tube experiments at $T \leqslant 1000$ K that were excluded in the general comparison.

The characterization of the influence of the total diluent concentration is even more difficult (Fig. 5, bottom right panel, left part). When more than one diluent was present, the sum of the diluent concentrations was considered. In experiments with high diluent concentrations ($X_{dil} > 0.95$), Ar was always used as the diluent, while in those with a low overall diluent concentration ($X_{dil} \leqslant 0.60$) the diluent was nitrogen. In the three medium intervals, diluents are mixed. Clear trends are not visible; the weak performance of all mechanisms at diluent intervals of 0.60–0.80 and 0.90–0.95 might be due to a superposition of other effects.

As opposed to the classic shock tube experiments, recorded pressure changes are available for the rapid compression machine (RCM) experiments and could be used in the simulations. However, none of the mechanisms were able to reproduce the low-temperature RCM measurements in a very satisfactory manner. Some mechanisms cannot reproduce any of the RCM experiments (Zsély-2005, Rasmussen-2008, Ahmed-2007, Sun-2007, SanDiego-2011). Their average error function values are magnitudes higher than those of the best mechanisms for RCM data and are therefore not discussed in detail. Yet, other mechanisms simulate these experiments well at temperatures above 960 K, especially Hong-2011, the Galway mechanisms (Kéromnès-2013, NUIG-NGM-2010 and ÓConaire-2004) and CRECK-2012 (see Fig. 5, top left panel, right part).

As opposed to the results of the shock tube simulations, a trend for the RCM data towards increasing error function values at higher pressure can be observed for all mechanisms (Fig. 5, bottom left panel, right part). Hong-2011 and Kéromnès-2013 are relatively insensitive to pressure variations and their performance at high pressures (26.5–32.3 atm and 49.3–69.1 atm) is very good.

For the dependencies on the equivalence ratio and the diluent concentration only three intervals could be distinguished. It can be seen that most reasonable-to-well performing mechanisms have the highest E values at low equivalence ratios ($\varphi = 0.35$) and low diluent concentrations ($X_{dil} = 0.65$–0.69), with exceptions in both cases.

In Fig. 2, top left panel, the correlation of the mechanisms with respect to the absolute deviation values are shown for shock tube simulations. Correlations of Dagaut-2003 (#19) with the other mechanisms are indicated with diagonal stripes as the number of simulations performed differed from the other mechanisms. It can be seen that most of the overall best-performing mechanisms (low identifying numbers) tend to be more strongly correlated with each other than with the weakly performing ones. The two mechanisms that perform the weakest for shock tube data, GRI3.0-1999 (#14) and Ahmed-2007 (#17) behave similarly to each other ($C = 0.909$) and behave considerably differently from almost all other mechanisms. The mechanism pairs

Davis-2005/USC-2007-II (#9/#11, $C = 0.999$), SaxenaWilliams-2006/SanDiego-2011 (#8/#13, $C = 0.981$) and NUIG-NGM-2010/ÓConaire-2004 (#2/#3, $C = 0.981$) were published by the same groups and have the highest correlated results for shock tube ignition delay times. For the shock tube data, it can be seen that the correlation values are a good indicator to describe similarities in the behavior of different mechanisms. On average, the RCM correlation values are lower than those for shock tube simulations (Fig. 2, top right panel). As for shock tube ignition delays, the best mechanisms (with the exception of Konnov-2008, #4) are also the higher correlated ones, again with the same pairs of mechanisms behaving practically identically: Davis-2005/USC-2007-II (#9/#11, $C = 1.00$) and NUIG-NGM-2010/ÓConaire-2004 (#2/#3, $C = 0.96$). While for shock tube simulations Saxena-Williams-2006 and SanDiego-2011 were very similar, they are not highly correlated for RCM data (#8/#13, $C = 0.74$).

6.2. Concentration–time profiles in flow reactors

As it can be seen in Fig. 6, for all mechanisms the prediction of O_2 profiles is much better than those for H_2 and H_2O recorded in flow reactors. In particular, the SaxenaWilliams-2006 and SanDiego-2011 mechanisms cannot accurately predict the concentrations of these species. Note that time shifting, which was performed based on 50% depletion of the fuel species H_2 separately for each mechanism, can have a strong influence on the agreement for the other species. This is the reason why flow reactor data are somewhat problematic for a comparison of different mechanisms with each other.

Figure 2 (bottom right panel) shows that the GRI3.0-1999 (#14), Dagaut-2003 (#19), Ahmed-2007 (#17), Starik-2009 (#10) and Konnov-2008 (#4) mechanisms are the least correlated for flow reactor simulations. Interestingly, all of these mechanisms are among the best performers in flow reactor simulations, except for GRI3.0-1999 and Konnov-2008. None of these mechanisms are particularly highly correlated amongst each other, except #17/#19. For the development of the Ahmed-2007 mechanism, flow reactor data for n-heptane were used for model validation [40], while for Dagaut-2003 their own flow reactor measurements

[26] were utilized. The mechanisms ranked 1st and 2nd for flow reactors, Starik-2009 and Burke-2012, were both validated against the H_2 flow reactor data of Mueller et al. [41], which explains the good flow reactor performance of these mechanisms.

6.3. Outlet concentrations in jet-stirred reactors (JSRs)

Figure 7 shows the performance of the mechanisms in reproducing outlet concentrations for various species measured in a jet-stirred reactor [42]. It can be seen that the prediction of H_2O concentrations is worse than for H_2 and O_2. Overall, most mechanisms show a very similar performance. The result for the GRI3.0-1999 mechanism is slightly better than those of the other mechanisms. As previously described, 100 data points at $T < 1000$ K were excluded from the comparison due to wall effects and are not shown in Fig. 7. GRI3.0-1999 is the only mechanism that is capable of predicting low-temperature JSR measurements with an acceptable accuracy. For all other mechanisms, the error function values are about one order of magnitude higher than at $T \geqslant 1000$ K. It is difficult to evaluate the remaining 49 data points at $T \geqslant 1000$ K with respect to the dependence of the error function values on specific operating conditions in a statistically meaningful manner. However, trends towards a worse prediction of JSR experiments by the mechanisms at high pressure ($p = 10$ atm) and very low equivalence ratio ($\varphi = 0.09$) can be observed. Further JSR studies would be valuable to confirm these trends and to extend the range of operating conditions covered by the measurements. With respect to the absolute deviations, all mechanisms behave almost identically for JSR simulations. Correlation values of most mechanisms are $C \geqslant 0.99$. They were therefore omitted in Fig. 2.

6.4. Flame velocity measurements

Unlike in the ignition delay and JSR measurements, He was used as the bath gas in several flame velocity determinations. Only mechanisms having He as the bath gas were used for the simulation of these experiments, which represent about one sixth of all flame velocity measurements. The general comparison of the mechanisms with respect to flame velocity data was carried out
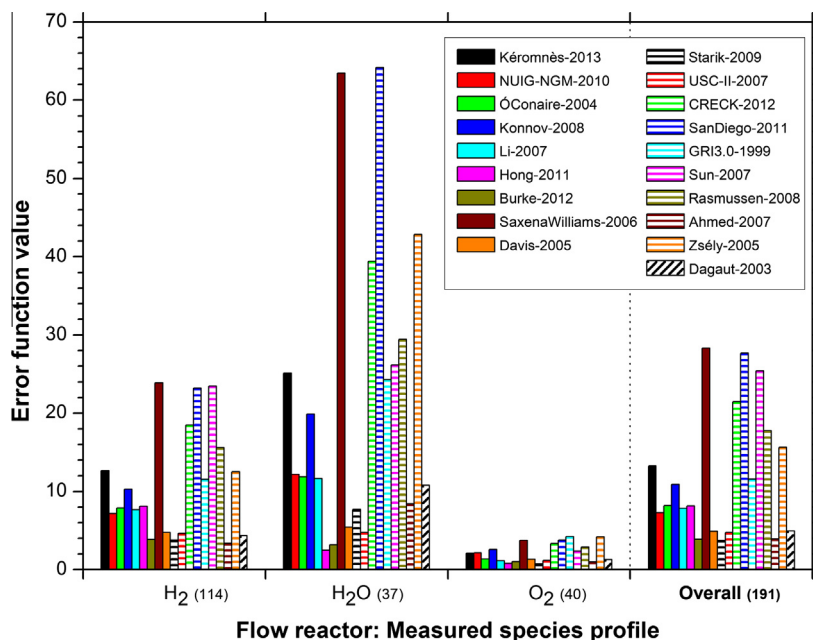

Fig. 6. Performance of the mechanisms with respect to flow reactor measurements for various measured profiles and overall.
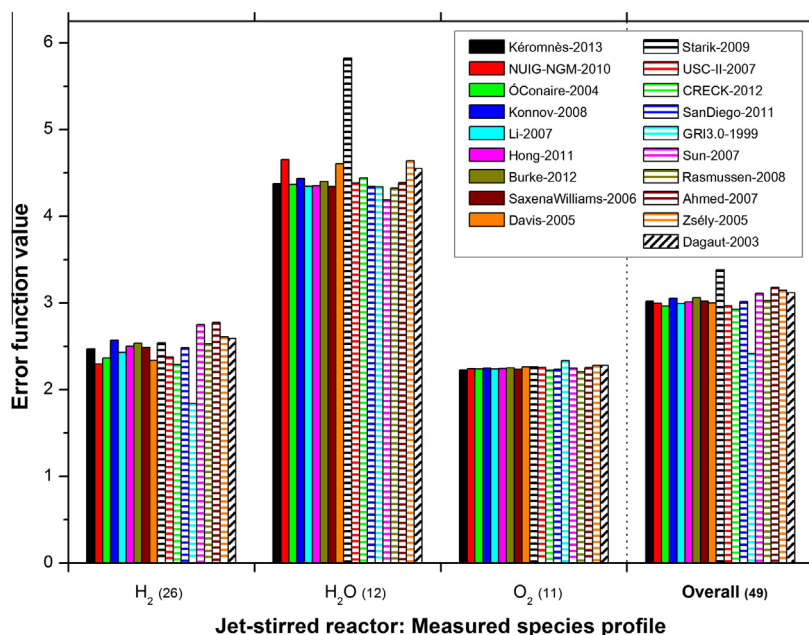
Fig. 7. Performance of the mechanisms with respect to JSR measurements for various measured profiles and overall.

without these experiments ("no He" in Fig. 10). A separate investigation for the measurements containing He as a diluent was performed for the 10 mechanisms in which this species was defined ("He").

Reproductions of the flame velocity measurements are much more similar than for ignition delay times. The best mechanisms are Kéromnès-2013, Burke-2012, Konnov-2008, Li-2007, Davis-2005 and SanDiego-2011 see Fig. 8, "Overall"). In contrast, CRECK-2012, Rasmussen-2008 and GRI3.0-1999 are the least able to reproduce the measured flame velocity results. Relative to their weak performance for ignition delay times, SanDiego-2011 and Ahmed-2007 are much more capable of predicting flame velocities. The opposite trend applied to Hong-2011.

Compared to ignition delay times, the mechanisms generally correlate much more strongly for flame velocity data, with a very similar general trend that the better the performance, the higher the correlation values (Fig. 2, bottom left panel). The highest correlated pairs of mechanisms are the mal-performing SaxenaWilliams-2006/SanDiego-2011 (#8/#13, $C = 0.999$) and the well performing Galway mechanisms NUIG-NGM-2010/ÓConaire-2004 (#2/#3, $C = 0.987$). Overall, the mechanisms of Sun-2007 (#15), Ahmed-2007 (#17) and CRECK-2012 (#12) are the least correlated compared to the others.

Flame velocities were measured using four different methods. These are the flame cone method (FCM) (see e.g., [43] for a discussion of this method), the outwardly propagating spherical flame method (OPF) [44], the counterflow twin-flame technique (CTF) [45], and the heat flux burner method (HFM) [46]. Figure 8 demonstrates that the reproduction of the experimental data measured by these methods improves in the order above. Experimental data measured with the flame cone method are especially poorly reproduced.

Figure 9 investigates the performance of the mechanisms to flame velocity measurements according to the type of the bath gas or diluent mixture. Interestingly, large deviations are observed when the performance of one mechanism for experiments in different diluents is studied as well as when all mechanisms are compared with respect to one type of diluent. However, for the

majority of the mechanisms, experiments using He-containing mixtures tend to be better-predicted than those with N_2-containing mixtures. With the exception of CRECK-2012 and Rasmussen-2008, measurements in mixtures of N_2 and Ar are predicted well; however, the sampling base was small for these mixtures. Some mechanisms (e.g., Davis-2005, Li-2007, Konnov-2008, SanDiego-2011, Burke-2012, Kéromnès-2013) are appropriate for most types of bath gases, while other mechanisms (like NUIG-NGM-2010, Hong-2011, CRECK-2012) perform well for one type of bath gas and poorly for another.

Figure 10 shows a comparison of the mechanisms with respect to temperature, pressure, equivalence ratio and total diluent concentration. Each of the four panels of the figure are sub-divided into two parts, of which the left part shows a general comparison for all data points except those measured in He and the right part shows the results for He-diluted experiments only, i.e., either in pure He, N_2/He or He/H_2O.

Only a few of the flame velocities were measured at conditions above room temperature (365 K in N_2, 393 K in He and He/H_2O). Meaningful conclusions on the influence of the pre-heat temperature on the mechanism performance cannot be drawn based on these scarce data. However, it is interesting that all mechanisms could reproduce the 21 measurement points at $T = 365$ K quite well. For the He-diluted experiments, some mechanisms perform well at $T = 393$ K (Davis-2005, Burke-2012, NUIG-NGM-2010, CRECK-2012), while others are less accurate (Starik-2009, Kéromnès-2013) at these conditions.

Figure 10 (bottom left panel) shows the general trend that the error of reproduction slightly decreases with increasing pressure, with the exception of a distinct maximum at intermediate pressure (except for Ahmed-2007). Most mechanisms, like those of Burke-2012 and Kéromnès-2013, show a uniform low deviation over all ranges of pressure. The worst performing mechanisms here are GRI3.0-1999 and Starik-2009, since their errors significantly increase at high pressures. For He-containing flames, pressure dependence was not observed for any of the mechanisms.

In contrast to ignition delay time data, a clear trend towards a decreasing performance with increasing equivalence ratio can be
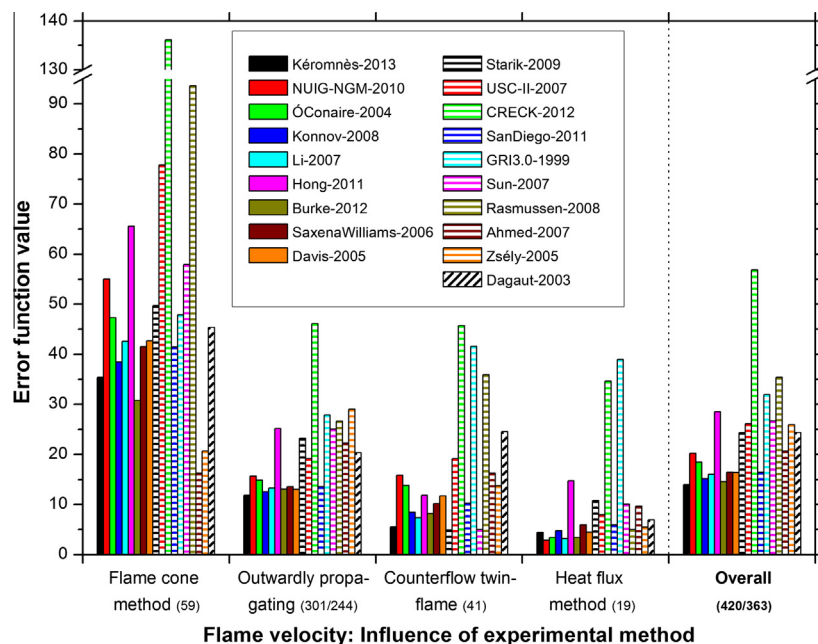
Fig. 8. Error of the reproduction of the flame velocity according to the type of experimental facility and overall.
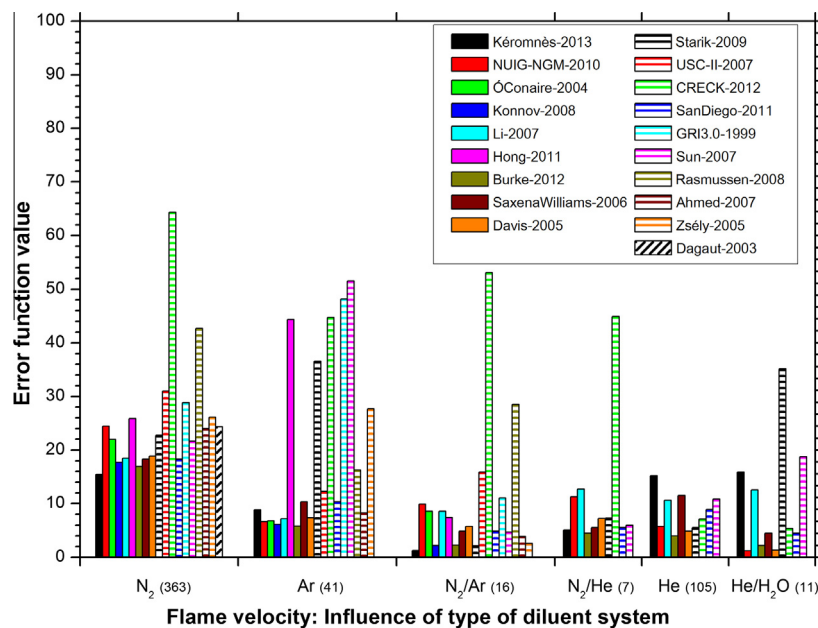

Fig. 9. Error of the reproduction of the flame velocity according to the type of diluent system.

observed for non-He flames, with the exception of CRECK-2012 (Fig. 10, top right panel). Interestingly, this trend does not apply to He-containing flames. For those flames, most of the mechanisms (except Sun-2007 and Starik-2009) have slightly higher E values at stoichiometric conditions. A trend towards lower error function values at higher diluent concentrations is displayed in Fig. 10, bottom right panel. Again, a similar relationship was not observed in He flames.

The mechanisms correlate more strongly for the results of flame simulations than for ignition delay times (Fig. 2), especially the overall best mechanisms #1 to #8 and #13. All correlation values among these mechanisms are higher than 0.90. The most weakly correlated mechanisms are Sun-2007 (#15), Ahmed-2007 (#17)

and CRECK-2012 (#12) with an overall minimum of $C = 0.40$ (#12/#17).

6.5. Considering all types of measurements simultaneously

In Fig. 11 the average absolute deviations for all mechanisms with respect to ignition delay time, JSR and flow reactor concentration data and flame velocity measurements are shown. Also, percentage values are given to describe the share of positive or negative deviation values for each mechanism. They should be read together with the actual average deviation to avoid misinterpretation. Note that a few data points with high D_{ij} values can heavily influence the average of a subset and its sign. The reader may notice
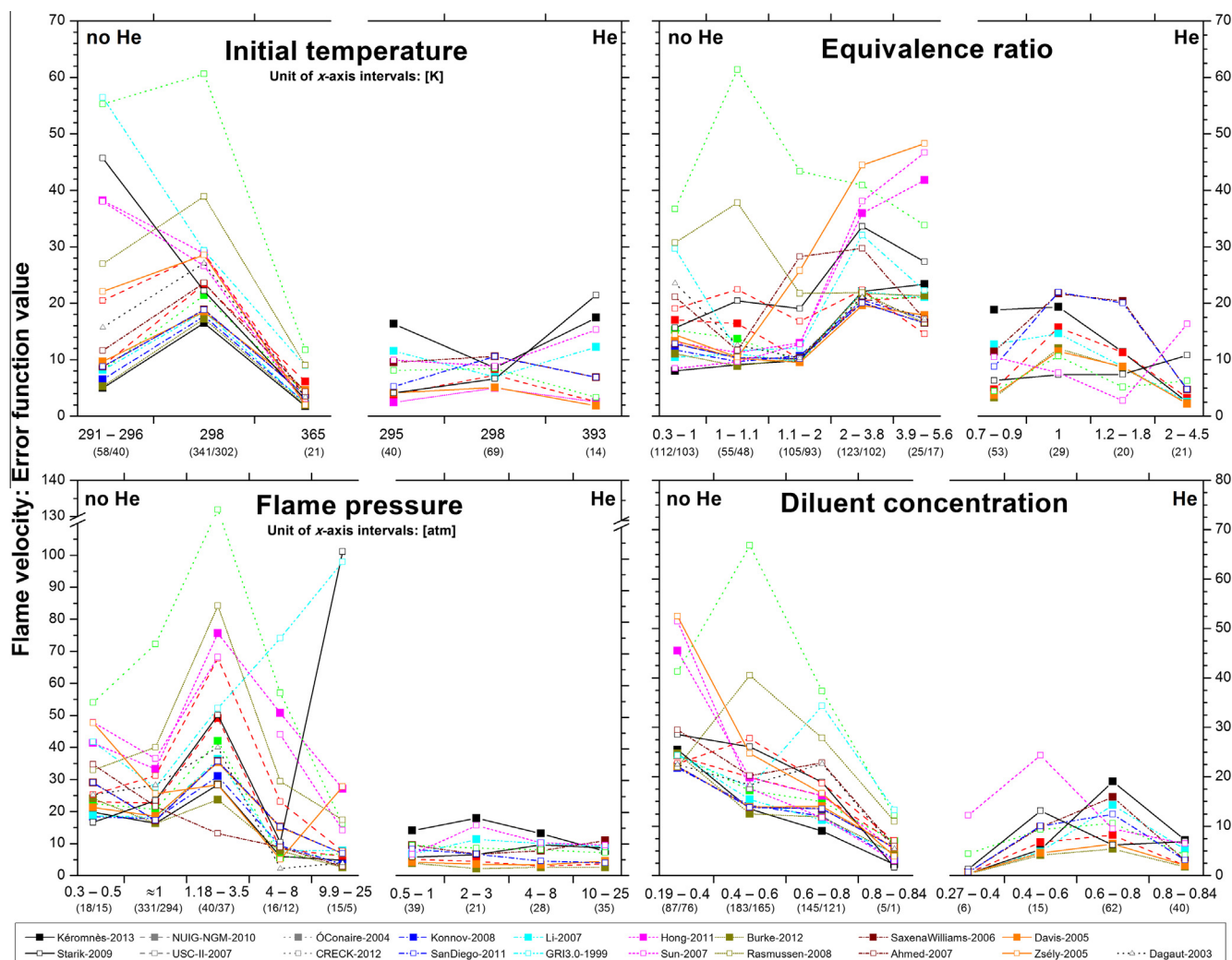
Fig. 10. Performance of the mechanisms for various ranges of temperature, pressure and equivalence ratio and diluent concentration with respect to flame velocities. Each plot shows the results for He-free mixtures (left part) and He-containing mixtures (right part). Note that those mechanisms without He as a defined species do not appear in the right part of the plots.

the orders of magnitude difference of the absolute deviation values among the types of measurements. All mechanisms tend towards over-prediction of the ignition delay time measurements, except Zsély-2005. Although there is a correlation between the good performance of a mechanism and the extent of over-prediction, the mechanism with an average absolute deviation of almost zero (Konnov-2008) is not the very best mechanism for ignition delay times (ranked 6th). This might be due to the averaging where positive and negative deviations might cancel out each other.

Flame velocities in He-free flames were under-predicted by all mechanisms except for Ahmed-2007 and Zsély-2005. The five worst performing mechanisms for hydrogen flames according to the descending error function values are CRECK-2012, Rasmussen-2008, GRI3.0-1999, Hong-2011, and USC-2007-II. These mechanisms also have the highest negative average absolute deviation values (order: CRECK-2012, Rasmussen-2008, Hong-2011, USC-2007-II, and GRI3.0-1999). For He-containing flames, the opposite trend applies: all mechanisms except CRECK-2012 tend to over-predict flame velocities (dark gray columns in Fig. 11). This could be a possible explanation for the diverging trends for He-containing flames observed in Fig. 10.

In the lower part of this figure, the average absolute deviations for JSR and flow reactor data are shown. For flow reactor data, a distinction between consumed species (H_2/O_2, over- or under-predicted

depending on the mechanism) and produced species (H_2O, under-predicted by the majority of mechanisms) was necessary due to different observed trends. For Hong-2011 it can be seen that despite a majority of 57% of the H_2/O_2 data points which were over-predicted by the simulations (i.e., too high concentrations), the average is below zero (and vice versa for H_2O concentrations). This can be an indication of outliers in the simulation results. The negative deviations from JSR experiments are, like the error function values, very similar for all mechanisms.

A special feature of the Li-2007 and Burke-2012 mechanisms (both published by the Princeton group) is the use of a different parameterization and third-body collision efficiency factors for the reaction $H + O_2(+M) = HO_2(+M)$ depending on whether Ar/He or N_2 are used as the main bath gas. In Fig. 12 the averaged simulation results of the two mechanisms handled according to the recommendation of the authors ("Adapted") are compared to an exclusive usage of one of the parameterizations ("All Ar/He" or "All N_2"). The RCM experiments of Kéromnès et al. [4] were carried out in a bath gas with an equal share of Ar and N_2, a special case that was not discussed by the authors of Li-2007 and Burke-2012. For these cases, the Ar/He parameterization was chosen as it generally performed better than the N_2 parameterization. Even though this option was meant to be applied for simulations with mixtures containing Ar and He, it of course handles the species
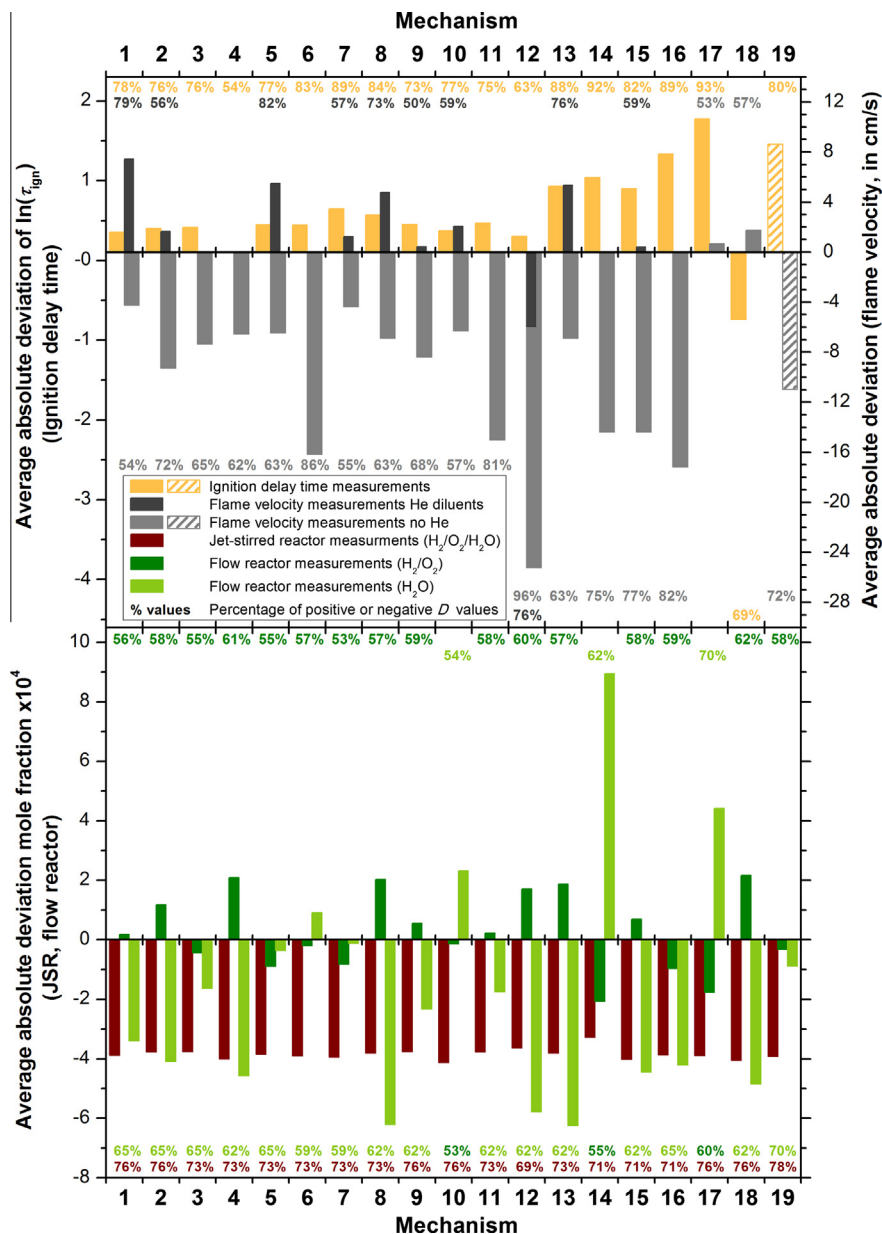
Fig. 11. Average absolute deviation values of all mechanisms for ignition delay times and flame velocities (top), and JSR and flow reactor concentrations depending on the measured species profile (bottom). The columns are striped when the number of data points differs from the other mechanisms (for Dagaut-2003, #19: all diluents except Ar and He).

N_2 too, and performs reasonably well. The average error function values are the lowest in the "Adapted" cases, which confirms the purpose of the procedure. What is interesting is that for the Burke-2012 mechanism, the changes due to the use of different parameterizations are very small, whereas for Li-2007 these are considerably large. For flow reactor data, the difference for the Li-2007 is due to a bad prediction of H_2O concentrations by the mechanism with Ar/He parametrization, while the bad average performance of the mechanism with N_2 parametrization for ignition delay times can be traced back to some experiments in Ar dilution. In the Li-2007 mechanism the collision efficiencies for the species Ar and N_2 in the reaction $H + O_2(+M) = HO_2(+M)$ were not specified (i.e., always 1), hence, they are identical for the two parametrizations. Furthermore, the F_{Cent} parameter that controls the transition between the low- and high-pressure limit rate parameter values differs between the two parametrizations of this reaction in the Li-2007 mechanism (Ar/He: $\alpha(F_{Cent}) = 0.5$, N_2: 0.8).

6.6. Analysis of the sensitivity coefficients

It is clear from the previous sections that mechanisms can differ significantly in their behavior. In order to improve a chemical model it might be useful to have information about which parts of a mechanism are responsible for a good or bad performance regarding a certain type of experiments or a constrained range of experimental conditions. Furthermore, it can be interesting to have an estimate of how the reproduction of some measurements changes if a reaction rate coefficient is tuned to describe another experiment better. To investigate these relationships, sensitivity analyses were performed at the conditions of all measurement data. Five mechanisms (#1, #2, #8, #12, #15) were arbitrarily chosen which represent different levels of predictivity and contain the species He, which allows a comparison based on the complete set of data. For the calculation of the sensitivity coefficients S, a brute force method was applied by varying the pre-exponential Arrhenius
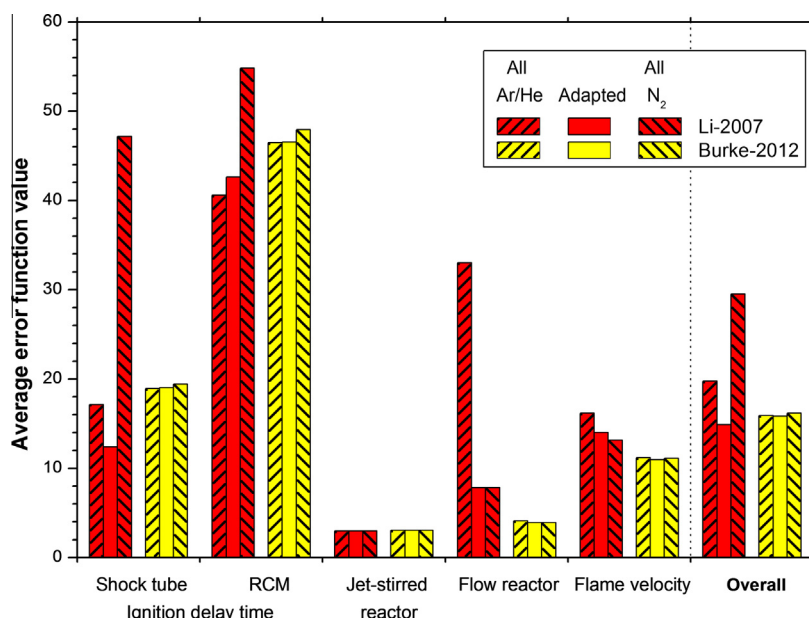
Fig. 12. Comparison of different parameterizations of the reaction H + O$_2$(+M) = HO$_2$(+M) in the Li-2007 and Burke-2012 mechanisms.

parameter A for all reactions by 10%, one-by-one. The S values were normalized and then scaled to the range of −1 to +1. This range was divided into 10 equidistant intervals and the relative frequency of the scaled sensitivity values for each of these intervals was then determined. It is very challenging to interpret and visualize the total of more than 230,000 S values generated from five mechanisms for each region of interest, which is the reason why only a few highlights of this analysis are presented here.

In Fig. 13, the distribution of the sensitivity coefficients for reaction H$_2$ + O = H + OH, which is one of the most important reactions in hydrogen combustion) with respect to all types of measurements is shown for four of these mechanisms. Note that a mechanism was excluded from the figures (e.g., mechanism Sun-2007, #15, from Fig. 13), if the respective reaction was not defined in a

trivial manner (e.g., in reverse direction or with 3rd body-dependent parametrizations of reactions). It is worth emphasizing that the sensitivity coefficients highly depend on the experimental conditions, which can differ largely among the types of measurements and facilities (e.g., shock tube and RCM measurements). In other words, different sub-chemistries can be important at various conditions across all types of measurements. If the experiments belonging to one measurement type were carried out in a narrow range of conditions (e.g., in JSRs and flow reactors), the corresponding distribution of the sensitivity coefficients cannot be directly compared to the ones of other measurement types (e.g., flames). In light of these limitations, it can be seen that H$_2$ + O = H + OH seems not to be very important at the conditions of RCM experiments and of jet-stirred reactors for the consumed species, while
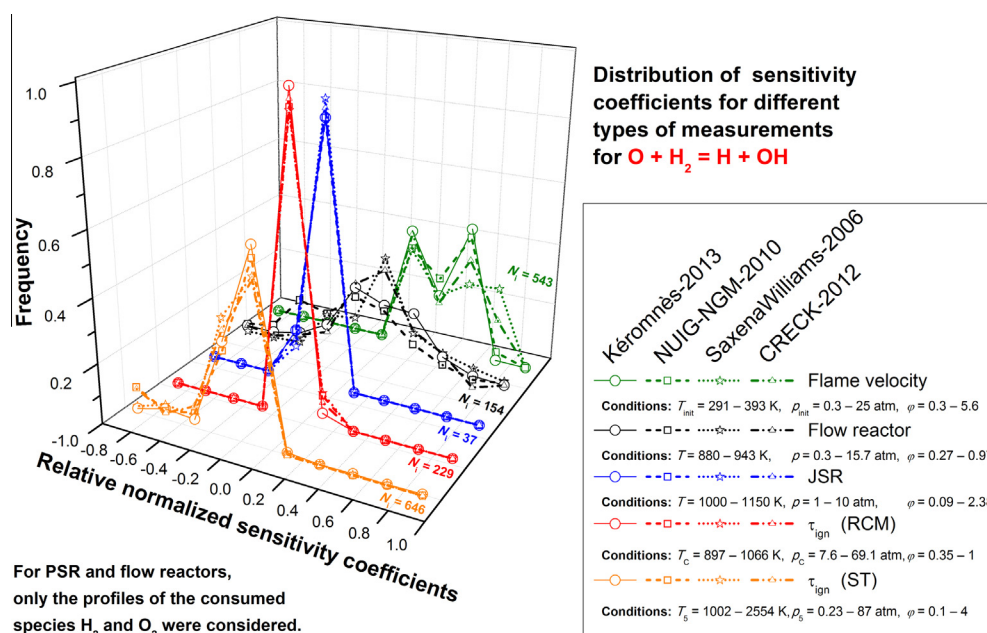

Fig. 13. Frequencies of sensitivity coefficients of reaction H$_2$ + O = H + OH for various types of measurements. The sensitivity analysis for this reaction was performed with the mechanisms Kéromnès-2013, NUIG-NGM-2010, SaxenaWilliams-2006 and CRECK-2012. The experimental conditions are given for each type of measurement.
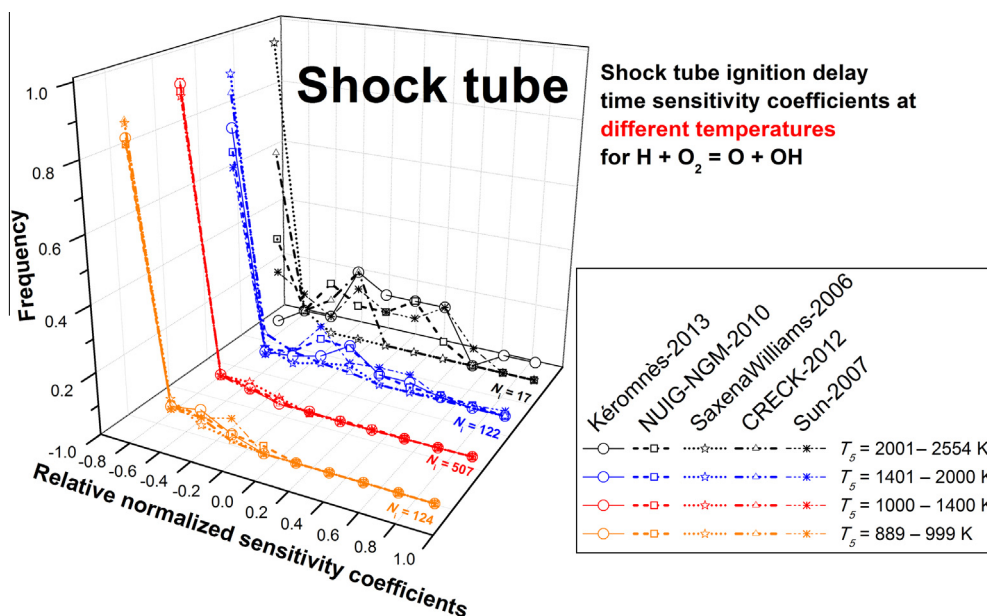
Fig. 14. Frequencies of sensitivity coefficients of reaction H + O₂ = O + OH for various ranges of temperature behind the reflected shock wave in shock tube experiments. The sensitivity analysis for this reaction was performed with the mechanisms Kéromnès-2013, NUIG-NGM-2010, SaxenaWilliams-2006, CRECK-2012 and Sun-2007.
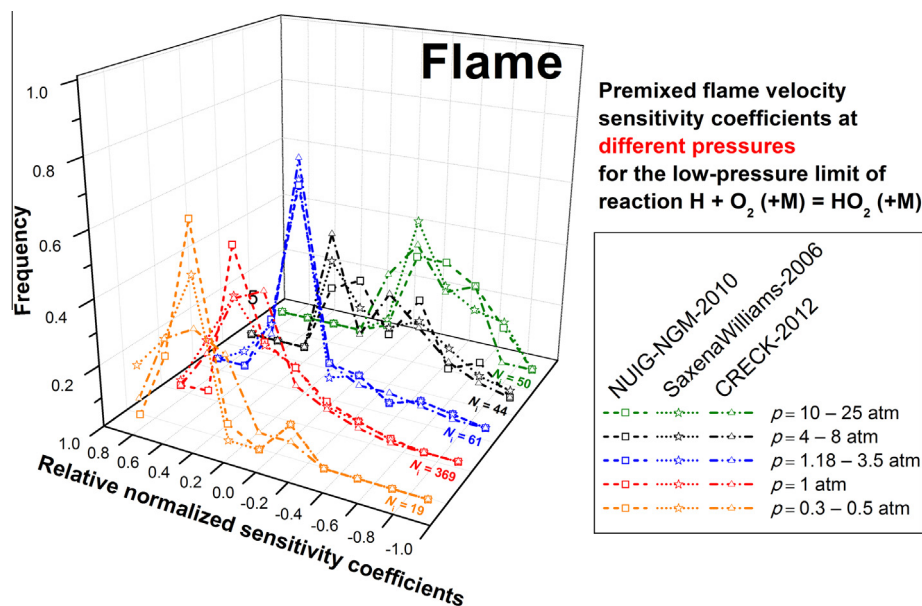

Fig. 15. Frequencies of sensitivity coefficients of the low-pressure limit of reaction H + O₂(+M) = HO₂(+M) for various ranges of flame pressure. The sensitivity analysis for this reaction was performed with the mechanisms NUIG-NGM-2010, SaxenaWilliams-2006 and CRECK-2012. The bottom axis was inverted to allow for a better comparison to Fig. 15 (high frequency on the left side corresponds to promoting the reactivity of the system).

it promotes reactivity in at the conditions of flames (positive S values, i.e., flame propagation becomes faster) and shock tubes (negative S values, i.e., ignition delays become shorter). For flow reactor data, the four mechanisms show a wide distribution of the sensitivity coefficients.

The influence of the crucial reaction H + O₂ = O + OH on shock tube ignition delay times depending on the temperature behind the reflected shock wave is shown in Fig. 14. Even though this reaction generally tends to increase reactivity, a decreasing relative importance of this reaction with increasing temperature can be observed (except for SaxenaWilliams-2006). This gives an indication that high-temperature chemistry in shock tubes is also dominated by other reactions, particularly the chain-terminating step

H + OH + M = H₂O + M (see Table G1 in the Supplemental material). Figure 15 illustrates that an increase of the rate coefficient of the low-pressure limit of reaction H + O₂(+M) = HO₂(+M) has a promoting effect on the reactivity in flames at low pressure, but tends to inhibit reactivity at higher pressures. A clear shift of the peaks of the histograms can be observed.

Generally the distributions of sensitivity coefficients are similar for all investigated mechanisms, although frequencies may vary at certain conditions. Hence, the information obtained from the sensitivity analysis of one mechanism can be used at least qualitatively for the improvement of other mechanisms. This statement is valid as long as the predominant reaction pathways are identical in the mechanisms, which is the case for hydrogen combustion.
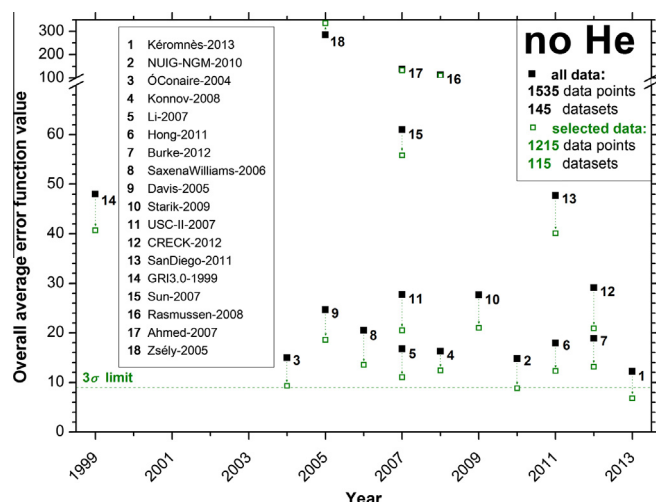
Fig. 16. Overall performance of the mechanisms taking into account all experimental data (black full square) and only those that were reproduced by at least one mechanism within $E_i \leqslant 9$ (green open square) *vs.* year of publication. All diluents except He.

The detailed sensitivity coefficients for the conditions explored in Figs. 5–7 and 10 are given in Tables G1–G5 of the Supplemental material exemplary for the Kéromnès-2013 mechanism.

By being able to access and examine a large amount of sensitivity data in connection with the results of the investigation of the performance of a mechanism, it becomes possible for model developers to identify those reactions which have to be revisited in their mechanisms in order to achieve a better agreement with the measured data. A possible further outcome of this investigation can be to identify less well-understood, but still important reactions in the hydrogen system that require additional theoretical calculations or experimental determinations of the reaction rate coefficients.

7. Conclusions

The accurate description of the combustion of hydrogen is important from both scientific and industrially applied points of view. Several excellent reviews were recently published [4,9–12], which discussed new developments in this field. However, a comprehensive investigation and comparison of all recent hydrogen combustion mechanisms has not been published.

Using black full squares, Fig. 16 shows the performance of each mechanism, tested against the experimental data in He-free diluent systems. The meaning of the open green square symbols will be discussed at the end of this chapter and should be disregarded for the time being. This figure indicates that the performance of the Kéromnès-2013 mechanism is currently the best, but several other mechanisms, such as the NUIG-NGM-2010, ÓConaire-2004, Konnov-2008 and Li-2007 mechanisms have similarly good overall performance. However, it is not advised to use the mechanisms of Zsély-2005, Ahmed-2007, Rasmussen-2008, Sun-2007, SanDiego-2011 and Dagaut-2003 for hydrogen combustion simulations, particularly not for RCM simulations. On the other hand, two of these mechanisms perform well for flow reactors (Ahmed-2007 and Dagaut-2003), which makes them a good choice for this type of measurements. The reason for some of the bad performances of certain mechanisms is due to lack of validation for certain types of measurements (e.g., flames for Rasmussen-2008). However, the majority of users of reaction mechanisms are interested in robust, general-purpose mechanisms, which is why the above recommendations are useful from the authors' viewpoint.

The best mechanisms for the reproduction of ignition delay times, flow reactor concentration profiles, JSR experiments and

flame velocity measurements are Kéromnès-2013, Starik-2009, GRI3.0-1999 and Kéromnès-2013, respectively. Several mechanisms do not work properly at high/low ranges of equivalence ratio, temperature, and pressure. Results, such as those presented in Figs. 2–12, may help in the selection of a mechanism best suited for the simulation of experiments at specific initial conditions. An analysis of sensitivity coefficients was carried out, yielding to some interesting conclusions for model developers.

Figure S1 in the Supplemental material shows the overall comparison results in the same manner, except that all data points (i.e., also the measurements in He-containing diluents) are covered. The trends and absolute values of E are similar to those in Fig. 16.

One major goal of this work was to identify regions of operating conditions that require further attention in mechanism development. On the other hand, measurements carried out at certain conditions might be accompanied by large experimental uncertainties or the measurement itself was problematic, e.g., due to assumptions in the interpretation of a measured signal. Although some measurements were excluded in the first place (e.g., Snyder et al. [47]), others that were retained in the comparison may still contain systematic errors. A possible approach to identify systematic errors is to optimize the rate parameters, thermochemical data and transport data of an appropriate mechanism on a single dataset only and excluding the datasets whose E_i values could not be lowered to values close to unity. Another approach is to exclude those measurements that none of the mechanisms could reproduce within a predefined uncertainty, e.g., 3σ, which is equivalent to an error function value of $E_i \leqslant 9$. In Tables A–E of the Supplemental material the experimental datasets that do not fulfill the $E_i \leqslant 9$ criterion are shaded gray. This procedure does not imply that these measurements necessarily have large systematic errors as it is possible that none of the mechanisms are capable of modeling the chemistry at a specific condition in an accurate manner. However, these measurements should be treated with care in future model development and should potentially receive further attention. In total, 31 of the 169 datasets in the overall comparison were excluded in a second round of comparison. By excluding the corresponding 330 data points, the performance of almost all mechanisms improved significantly, as it is indicated by green open square symbols in Fig. 16. An average agreement of two mechanisms (Kéromnès-2013 and NUIG-NGM-2010) with the experimental data of less or equal than 3σ can be achieved this way.

Acknowledgments

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.combustflame.2014.03.006.

References

[1] T. Turányi, T. Nagy, I.G. Zsély, M. Cserháti, T. Varga, B.T. Szabó, I. Sedyó, P.T. Kiss, A. Zempléni, H.J. Curran, Int. J. Chem. Kinet. 44 (2012) 284–302.
[2] I.G. Zsély, T. Varga, T. Nagy, M. Cserháti, T. Turányi, S. Peukert, M. Braun-Unkhoff, C. Naumann, U. Riedel, Energy 43 (2012) 85–93.
[3] T. Varga, I. G. Zsély, T. Turányi, T. Bentz, M. Olzmann, Int. J. Chem. Kinet., 2013. http://dx.doi.org/10.1002/kin.20829.

[4] A. Kéromnès, W.K. Metcalfe, K.A. Heufer, N. Donohoe, A.K. Das, C.-J. Sung, J. Herzler, C. Naumann, P. Griebel, O. Mathieu, M.C. Krejci, E.L. Petersen, W.J. Pitz, H.J. Curran, Combust. Flame 160 (6) (2013) 995–1011.
[5] M. Tamura, P.A. Berg, J.E. Harrington, J. Luque, J.B. Jeffries, G.P. Smith, D.R. Crosley, Combust. Flame 114 (3–4) (1998) 502–514.
[6] T. Kathrotia, M. Fikri, M. Bozkurt, M. Hartmann, U. Riedel, C. Schulz, Combust. Flame 157 (7) (2010) 1261–1273.
[7] J. Herzler, C. Naumann, Proc. Combust. Inst. 32 (1) (2009) 213–220.
[8] J. Li, Z. Zhao, A. Kazakov, F.L. Dryer, Int. J. Chem. Kinet. 36 (2004) 566–576.
[9] M. ÓConaire, H.J. Curran, J.M. Simmie, W.J. Pitz, C.K. Westbrook, Int. J. Chem. Kinet. 36 (11) (2004) 603–622.
[10] A.A. Konnov, Combust. Flame 152 (4) (2008) 507–528.
[11] M.P. Burke, M. Chaos, Y. Ju, F.L. Dryer, S.J. Klippenstein, Int. J. Chem. Kinet. 44 (7) (2012) 444–474.
[12] Z. Hong, D.F. Davidson, R.K. Hanson, Combust. Flame 158 (2011) 633–644.
[13] P. Saxena, F.A. Williams, Combust. Flame 145 (2006) 316–323.
[14] S.G. Davis, A.V. Joshi, H. Wang, F. Egolfopoulos, Proc. Combust. Inst. 30 (2005) 1283–1292.
[15] A.M. Starik, N.S. Titova, A.S. Sharipov, V.E. Kozlov, Combust. Explo. Shock Waves 46 (5) (2010) 491–506.
[16] CRECK modeling Group, Hydrogen/CO Mechanism Version 1201. <http://creckmodeling.chem.polimi.it/kinetic.html/>.
[17] H. Sun, S.I. Yang, G. Jomaas, C.K. Law, Proc. Combust. Inst. 31 (2007) 439–446.
[18] C.L. Rasmussen, J. Hansen, P. Marshall, P. Glarborg, Int. J. Chem. Kinet. 40 (8) (2008) 454–480.
[19] I.G. Zsély, J. Zádor, T. Turányi, Proc. Combust. Inst. 30 (1) (2005) 1273–1281.
[20] D. Healy, D.M. Kalitan, C.J. Aul, E.L. Petersen, G. Bourque, H.J. Curran, Energy Fuel 24 (2010) 1521–1528.
[21] J. Li, Z. Zhao, A. Kazakov, M. Chaos, F.L. Dryer, J.J.J. Scire, Int. J. Chem. Kinet. 39 (2007) 109–136.
[22] H. Wang, X. You, A.V. Joshi, S.G. Davis, A. Laskin, F. Egolfopoulos, C.K. Law, USC Mech Version II. High-temperature Combustion Reaction Model of H$_2$/CO/C$_1$–C$_4$ Compounds. <http://ignis.usc.edu/USC_Mech_II.htm/>.
[23] Mechanical and Aerospace Engineering (Combustion Research), University of California at San Diego: Chemical-Kinetic Mechanisms for Combustion Applications, San Diego Mechanism, version 2011-11-22. <http://combustion.ucsd.edu/>.
[24] G.P. Smith, D.M. Golden, M. Frenklach, N.W. Moriary, B. Eiteneer, M. Goldenberg, C.T. Bowman, R.K. Hanson, S. Song, W.C. Gardiner, V.V. Lissianski, Z. Qin, GRI-Mech 3.0. <http://www.me.berkeley.edu/gri_mech/>.
[25] S.S. Ahmed, F. Mauß, G. Moréac, T. Zeuch, Phys. Chem. Chem. Phys. 9 (2007) 1107–1126.
[26] P. Dagaut, F. Lecomte, J. Mieritz, P. Glarborg, Int. J. Chem. Kinet. 35 (2003) 564–575.
[27] T. Turányi, MECHMOD v. 1.42: Program for the Transformation of Kinetic Mechanisms. <http://garfield.chem.elte.hu/Combustion/mechmod.htm/>.
[28] M. Frenklach, PrIMe Webpage. <http://www.primekinetics.org/>.
[29] R.J. Kee, F.M. Rupley, J.A. Miller, Sandia National Laboratories Report SAND89-8009B, 1989.
[30] A.E. Lutz, R.J. Kee, J.A. Miller, Sandia National Laboratories Report SAND87-8248, 1988.
[31] G.A. Pang, D.F. Davidson, R.K. Hanson, Proc. Combust. Inst. 32 (1) (2009) 181–188.
[32] M. Chaos, F.L. Dryer, Int. J. Chem. Kinet. 42 (3) (2010) 143–150.
[33] F.L. Dryer, M. Chaos, Combust. Flame 152 (2008) 293–299.
[34] M. Chaos, F.L. Dryer, Combust. Sci. Technol. 180 (6) (2008) 1053–1096.
[35] A.K. Das, C.-J. Sung, Y. Zhang, G. Mittal, Int. J. Hydrogen Energy 37 (8) (2012) 6901–6911.
[36] G. Mittal, C.J. Sung, R.A. Yetter, Int. J. Chem. Kinet. 38 (2006) 516–529.
[37] P. Glarborg, R.J. Kee, J.F. Grcar, J.A. Miller, Sandia National Laboratories Report SAND86-8209, 1986.
[38] R.J. Kee, J.F. Grcar, M.D. Smooke, J.A. Miller, Sandia National Laboratories Report SAND85-8240, 1985.
[39] D.L. Baulch, C.T. Bowman, C.J. Cobos, R.A. Cox, T. Just, J.A. Kerr, M.J. Pilling, D. Stocker, J. Troe, W. Tsang, R.W. Walker, J. Warnatz, J. Phys. Chem. Ref. Data 34 (3) (2005) 757–1397.
[40] C.V. Callahan, T.J. Held, F.L. Dryer, R. Minetti, M. Ribaucour, L.R. Sochet, T. Faravelli, P. Gaffuri, E. Rani, Proc. Combust. Inst. 26 (1) (1996) 739–746.
[41] M.A. Mueller, T.J. Kim, R.A. Yetter, F.L. Dryer, Int. J. Chem. Kinet. 31 (1999) 113–125.
[42] T. Le Cong, P. Dagaut, Energy Fuel 23 (1) (2009) 725–734.
[43] G.E. Andrews, D. Bradley, Combust. Flame 18 (1972) 133–153.
[44] D. Bradley, P.H. Gaskell, X.J. Gu, Combust. Flame 104 (1996) 176–198.
[45] C.K. Wu, C.K. Law, Proc. Combust. Inst. 20 (1984) 1941–1949.
[46] L.P.H. de Goey, A. van Maaren, R.M. Quax, Combust. Sci. Technol. 92 (1993) 201–207.
[47] A.D. Snyder, J. Robertson, D.L. Zanders, G.B. Skinner, Technical Report AFAPL-TR-65-93, 1965.